

# An Improved Forecast of Patchy Reionization Reconstruction with CMB

Meng Su<sup>1,\*</sup>, Amit P.S. Yadav<sup>2</sup>, Matthew McQuinn<sup>3</sup>, Jaiyul Yoo<sup>4,5</sup>, and Matias Zaldarriaga<sup>2</sup>

<sup>1</sup> *Institute for Theory and Computation, Harvard-Smithsonian Center for Astrophysics,  
60 Garden Street, MS-10, Cambridge, MA 02138*

<sup>2</sup> *School of Natural Sciences, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540*

<sup>3</sup> *Department of Astronomy, University of California, Berkeley, CA 94720*

<sup>4</sup> *Institute for Theoretical Physics, University of Zurich,  
Winterthurerstrasse 190, CH-8057 Zurich, Switzerland and*

<sup>5</sup> *Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720*

Inhomogeneous reionization gives rise to angular fluctuations in the Cosmic Microwave Background (CMB) optical depth  $\tau(\hat{n})$  to the last scattering surface, correlating different spherical harmonic modes and imprinting characteristic non-Gaussianity on CMB maps. Recently the minimum variance quadratic estimator  $\hat{\tau}(\hat{n})$  has been derived using this mode-coupling signal, and found that the optical depth fluctuations could be detected with  $(S/N)^2 \sim 100$  in futuristic experiments like CMBPol. We first demonstrate that the non-Gaussian signal from gravitational lensing of CMB is the dominant source of contamination for reconstructing inhomogeneous reionization signals, even with 98% of its contribution removed by delensing. We then construct unbiased estimators that simultaneously reconstruct inhomogeneous reionization signals  $\tau(\hat{n})$  and gravitational lensing potential  $\phi(\hat{n})$ . We apply our new unbiased estimators to future CMB experiment to assess the detectability of inhomogeneous reionization signals. With more physically motivated simulations of inhomogeneous reionizations that predict an order of magnitude smaller  $C_l^{\tau\tau}$  than previous studies, we show that a CMBPol-like experiment could achieve a marginal detection of inhomogeneous reionization,  $(S/N)^2 \sim \mathcal{O}(1)$  with this quadratic estimator to  $\sim \mathcal{O}(10)$  with the analogous maximum likelihood estimator.

PACS numbers:

## I. INTRODUCTION

Reionization marks the time in which the vast majority of the hydrogen in the Universe was ionized. When and how this process occurred is at present poorly constrained. Current data show that it must have ended by  $z \approx 6$  because at lower redshifts there was significant transmission in the Ly $\alpha$  forest [1]. In addition, the large-scale polarization anisotropies in the cosmic microwave background (CMB) constrain the mean redshift of reionization to have been  $z = 10.6 \pm 1.2$  [2].

It is believed that the first galaxies in the Universe produced the ionizing photons that ultimately ionized the intergalactic gas (e.g. [3]). The morphology of reionization and its duration depended on the nature, abundance, and clustering of the ionizing sources [4, 5]. There are several established ideas for how to better constrain the morphology and the redshifts over which it occurred. These include detecting the reionization-induced suppression and spatial modulation in the statistics of high-redshift Lyman- $\alpha$  emitting galaxies [6–8], studying H I Lyman- $\alpha$  damping wing absorption from the neutral gas during reionization in the afterglow spectra of high-redshift gamma ray bursts [6, 9, 10], and directly observing 21 cm emission from  $z > 6$  neutral hydrogen in the intergalactic medium (e.g., Furlanetto et al. [11]). This study concentrates on using a new technique, first proposed in Dvorkin

and Smith [12], that exploits the non-Gaussianities in the CMB sourced by reionization to study this process.

Inhomogeneous reionization produces several secondary anisotropies in the CMB. First, extra temperature (and, to a lesser extent, polarization) anisotropies are generated from peculiar motion of ionized regions during the entire reionization process [13–19]. These anisotropies are termed the kinetic Sunyaev-Zeldovich effect. Second, ionized bubbles scatter the local CMB temperature quadrupole, generating fluctuations in the polarization at *large scales* [15, 20]. Finally, the patchy nature of reionization would have resulted in the Thomson scattering optical depth to recombination,  $\tau(\hat{n})$ , depending on direction [15, 16, 21–23]. Such optical depth fluctuations act as a modulation effect on CMB fields by suppressing the primordial anisotropies with a factor of  $e^{-\tau(\hat{n})}$ , correlating different spherical harmonics. Information contained in  $\tau(\hat{n})$  fluctuations could potentially probe the duration of hydrogen reionization and the size of the ionized regions.

It is well known that gravitational lensing also imprints a non-Gaussian signature on the CMB. Minimum variance quadratic estimator has been introduced by using the coupled modes to reconstruct the projected lensing potential [24–27]. Recently, Dvorkin and Smith [12] followed similar technique and derived the minimum variance quadratic estimator for  $\tau(\hat{n})$ . Utilizing a toy model for reionization, they estimated that the patchy reionization signal could be detected with  $(S/N)^2 \sim 100$  for a CMBPol-like experiment, with beam full-width half-maximum (FWHM) of  $\Theta_{\text{FWHM}} = 4'$ , and noise sensitiv-

---

\*Electronic address: mengsu@cfa.harvard.edu

TABLE I: Description of various reionization models. Distances are in comoving units.

	$\zeta_7$	$d\zeta/dz$	$l_{\text{mfp}}$ (Mpc)	$\tau$	$L_{\text{box}}$ (Mpc)	$N_g$
A	10	0	$\infty$	0.063	200	256
B	10	18	$\infty$	0.112	200	256
C	30	0	10	0.090	200	256
D	20	16	10	0.115	200	256

ity  $\Delta_T = 1\mu\text{k-arcmin}$ . In this paper, we quantify the impact of lensing induced non-Gaussianities on the reconstruction of  $\tau(\hat{\mathbf{n}})$ . We show that lensing biases the reconstruction of  $\tau(\hat{\mathbf{n}})$ , and as a solution we construct an unbiased estimator for  $\tau(\hat{\mathbf{n}})$  in the presence of lensing.

The structure of this paper is as follows. Section II provides simple estimates for the size of  $\tau(\hat{\mathbf{n}})$  fluctuations, and it describes the cosmological reionization calculations used here to produce  $\tau(\hat{\mathbf{n}})$  maps. Section III derives the minimum variance quadratic estimator for  $\hat{\tau}(\hat{\mathbf{n}})$  in the flat sky approximation, and it quantifies the effect on lensing on the estimator. In section VI, we summarize our results and discuss the implications. In Appendix B, we describe in more detail our simulations to reconstruct  $\tau(\hat{\mathbf{n}})$  in the presence of lensing.

## II. INHOMOGENEOUS REIONIZATION

Patchy reionization produced a line-of-sight dependent optical depth that can be written as

$$\tau(\hat{\mathbf{n}}) = c \int \frac{a dz}{H(z)} \sigma_T \bar{n}_e(z) [1 + \delta_b(\hat{\mathbf{n}}, z) + \delta_x(\hat{\mathbf{n}}, z)], \quad (1)$$

where  $\sigma_T$  is the Thompson scattering optical depth,  $\bar{n}_e(z)$  is the average free-electron proper number density, and  $\delta_b$  and  $\delta_x$  are the over-densities in baryons and in the ionized fraction,  $x_i$ .

The angular power spectrum  $C_l^{\tau\tau}$  in the flat sky approximation can be related to the 3D ionization and density field using the Limber approximation:

$$C_l^{\tau\tau} = \int \frac{d\eta}{\eta^2} a^2 \sigma_T^2 \bar{n}_e(z)^2 \left[ P_{xx}(z, \frac{1}{\eta}) + 2 P_{x\delta}(z, \frac{1}{\eta}) + P_{\delta\delta}(z, \frac{1}{\eta}) \right], \quad (2)$$

where  $\eta$  is the conformal distance from the observer, and  $P_{XY}$  is the cross power spectrum of the over-density in  $X$  with the overdensity in  $Y$ . This power spectrum is weighted heavily to the highest redshifts where there was reionized gas. The kinetic Sunyaev-Zeldovich effect (kSZ) signal from reionization is predicted to be comparable to the kSZ signal after reionization. However, the kSZ weights by an additional  $v^2$  factor which goes roughly as the scale factor, result in its kernel peaking at lower redshifts [15]. This results in a large fraction of the kSZ coming from after reionization, we can safely neglect the low

redshift part, whereas we expect most of  $C_l^{\tau\tau}$  to originate from during reionization. It is also clear from equation (2) that  $C_l^{\tau\tau}$  from reionization increases approximately linearly with the duration of reionization for fixed mean redshift of reionization.

It is likely that reionization occurred in a patchy manner, with some regions being ionized early on in this process and others remaining neutral until the end, and with little gas at intermediate ionization states. This patchiness likely resulted in the ionization fluctuations dominating over other sources of fluctuation (i.e.,  $P_{xx} \gg P_{\delta\delta}$  on arcmin and larger scales [4]). Even without any knowledge of  $P_{xx}$  other than that reionization was patchy, there is an integral constraint on  $C_l^{\tau\tau}$  because if the ionization field is zeros and ones  $\int k^2 dk / (2\pi^2) P_{xx} = x_i^{-1} - 1$ , where  $x_i$  is the ionized fraction. Thus, fixing the reionization history and in the Limber approximation,  $\int l^2 dl C_l^{\tau\tau}$  is just a single number independent of morphology. This constraint shows that the larger the H II regions during reionization, the larger the fluctuations in  $\tau(\hat{\mathbf{n}})$ <sup>1</sup>.

To estimate  $C_l^{\tau\tau}$ , we compute Monte-Carlo realizations of reionization in two hundred comoving Mpc data cubes using the method developed in [28] for assigning the ionization state to boxes with realizations of the linear-theory cosmological density field. This method is based on the semi-analytic model for reionization in [4]. The distribution of ionized gas found in the [28] method is in excellent agreement with the results of detailed numerical simulations of reionization [28, 29]. Thus, we expect that the  $\tau$  field from this simulation will be more realistic than the analytic model used in the original study of [12]. Their model assumed a lognormal distribution of bubbles with a distribution that was independent of ionized fraction. In our calculations, the morphology of the bubbles is complicated and their sizes increase dramatically as  $x_i$  increases.

The method in [28] that we employ posits that the number of galaxies within a region sets its ionization state. Namely, a region is ionized if  $1 > \zeta f$ , where  $\zeta$  is a factor that encodes the efficiency that galaxies can ionize their surroundings, and  $f$  is the total fraction that has collapsed into halos with mass  $> m_{\text{min}}$ , where  $m_{\text{min}}$  is the minimum halo mass of the sources during reionization. A point in space is marked as ionized if this criterion is met for any smoothing scale centered around it (where the smoothing is done with a tophat in Fourier space filter).

Calculating  $f$  in detail requires high-resolution  $N$ -body simulations to resolve  $\sim 10^8 M_\odot$  halos – the smallest halos that were expected to form multiple generations of stars –, while still capturing scales much larger than the 10 comoving Mpc bubbles. Fortunately, extended Press-

<sup>1</sup> Since  $l^3 C_l \sim \text{const.}$ , if the peak in the bubble scale is at smaller  $l$  (i.e. larger bubbles), as  $l^2 C_l \sim l^{-1}$ , larger bubbles result in a higher peak (but at lower  $l$  i.e. larger fluctuations).

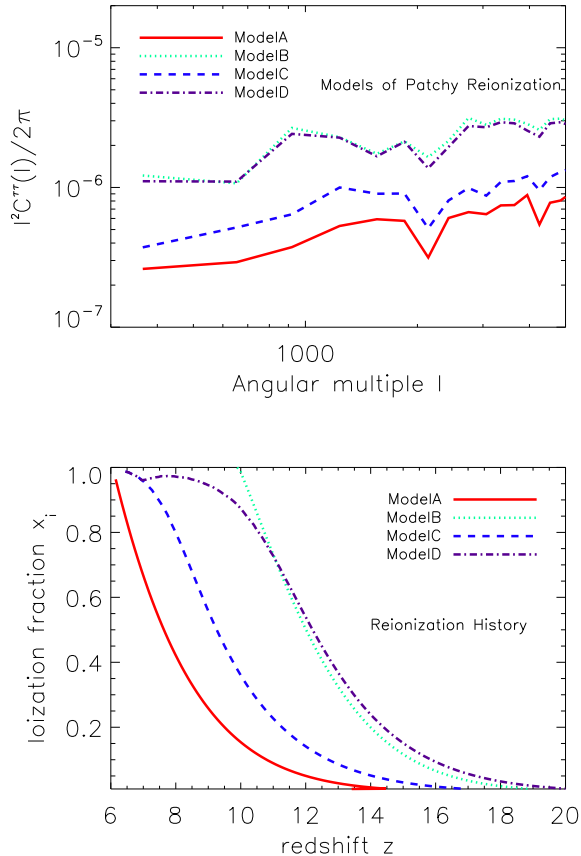


FIG. 1: *Upper panel:* Optical depth power spectrum from the different reionization models described in Table I. *Lower panel:* Corresponding average reionization history of these models. The mean optical depth of each model is consistent with the WMAP measurement ( $\tau = 0.088 \pm 0.015$ ).

Schechter theory provides a method to calculate  $f$  in a macroscopic region of size  $R$  and overdensity  $\delta_R$  in the simulations from just the linear density field [30, 31]. Therefore, we can quickly compute the ionization field from the linear density field and rather coarse resolution using just Fast Fourier Transforms. Our calculations take just minutes on a single CPU for the  $256^3$  grids used here.

There is significant uncertainty in the properties of the first sources and sinks that were responsible for reionization. All of our models assume that the ionizing luminosity is proportional to the collapsed mass in halos above  $10^8 M_\odot$  (approximately the minimum mass threshold where the gas can cool by atomic transitions and form dense structures). To explore the allowed parameter space of this process, we model reionization with 3 parameters: the ionizing efficiency of a halo  $\zeta_7$  at  $z = 7$ , its derivative with redshift  $d\zeta/dz$  (assumed to be independent of  $z$ ), and the mean free path of ionizing photons to be absorbed by an over-dense sink  $l_{\text{mfp}}$  within an ionized region. The first two parameters primarily affect

the duration of reionization while the later parameter primarily affects its morphology [5, 32].

In particular, bubbles that are larger than mean free path of ionizing photons have most of the photons produced within them absorbed by dense systems within the bubble rather than by diffuse gas at the bubble edge, preventing further growth [5, 32]. Thus, the parameter  $l_{\text{mfp}}$  is implemented by setting the maximum smoothing scale used to be  $l_{\text{mfp}}$ . We generate Monte-Carlo maps for 4 different reionization models, which are described in Table I. All of our models fall within  $2\sigma$  of the best fit WMAP  $\tau$  measurement of  $\tau = 0.088 \pm 0.015$  [2].

In Fig. 1, the top panel shows the optical depth fluctuation power spectrum of the different reionization models described in Table I. The corresponding reionization history of the four models are also shown in the bottom panel of Fig. 1. Surprisingly, the spectrum of all these models is not significantly different: All the models scales as  $l^2 C_l^{\tau} / 2\pi \approx \text{constant}$  for  $200 \lesssim l \lesssim 10000$ . However, the amplitude varies between  $\sim 10^{-6} - 10^{-7}$ , owing to the different reionization histories. An amplitude of  $10^{-6}$  is still an order of magnitude smaller than the signal considered in the previous work of [12]. It is possible that reionization is more extended than in our models. We note that the amplitude of  $C_l^{\tau}$  is proportional to the duration of reionization.<sup>2</sup>

### III. STANDARD QUADRATIC ESTIMATOR OF PATCHY REIONIZATION FROM THE CMB

The observed CMB temperature and polarization Stokes parameters in the presence of inhomogeneous screening caused by patchy ionized regions are

$$\begin{aligned} T(\hat{n}) &= e^{-\delta\tau(\hat{n})} \tilde{T}(\hat{n}), \\ (Q \pm iU)(\hat{n}) &= e^{-\delta\tau(\hat{n})} (\tilde{Q} \pm i\tilde{U})(\hat{n}), \end{aligned} \quad (3)$$

where tildes signify the CMB Stokes parameters for a uniform reionization history with constant factor  $e^{-\tau}$  spatially modulating the observed CMB fields. We take  $\tau$  as the mean of optical depth and  $\delta\tau(\hat{n})$  as the line of sight dependent optical depth fluctuation field. We work

<sup>2</sup> Recently it was shown that the velocity difference between the baryons and dark matter that is imparted up until recombination and decays away thereafter, can suppress the formation and baryonic accretion of the  $\lesssim 10^6 M_\odot$  halos that harbor the first stars [33, 34]. The standard paradigm is that these halos did not contribute significantly to reionization [11], but they may have ionized the intergalactic medium fractionally. Different regions in the Universe have different velocity offsets, with the coherence length of this difference being hundreds of Mpc. Even if these first stars just fractionally ionized the Universe, this large-scale modulation of the velocity difference could lead to larger fluctuations in  $\tau(\hat{n})$  (and peaking at  $l \sim 100$ ) than in the models we have considered. Thus, we point out that there remains the possibility of generating a larger signal than in the models considered here.

TABLE II: Minimum variance filters for optical depth estimator  $\hat{\tau}(\hat{\mathbf{n}})$  and lensing potential estimator  $\hat{\phi}(\hat{\mathbf{n}})$ 

$XX'$	$f_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2)$	$f_{XX'}^{lens}(\mathbf{l}_1, \mathbf{l}_2)$
$TT$	$C_{l_1}^{TT} + C_{l_2}^{TT}$	$C_{l_1}^{TT}(\mathbf{l} \cdot \mathbf{l}_1) + C_{l_2}^{TT}(\mathbf{l} \cdot \mathbf{l}_2)$
$TE$	$\tilde{C}_{l_1}^{TE} \cos 2(\varphi_{l_1} - \varphi_{l_2}) + C_{l_2}^{TE}$	$\tilde{C}_{l_1}^{TE}(\mathbf{l} \cdot \mathbf{l}_1) \cos 2(\varphi_{l_1} - \varphi_{l_2}) + C_{l_2}^{TE}(\mathbf{l} \cdot \mathbf{l}_2)$
$TB$	$\tilde{C}_{l_1}^{TE} \sin 2(\varphi_{l_1} - \varphi_{l_2})$	$\tilde{C}_{l_1}^{TE}(\mathbf{l} \cdot \mathbf{l}_1) \sin 2(\varphi_{l_1} - \varphi_{l_2})$
$EE$	$[\tilde{C}_{l_1}^{EE} + \tilde{C}_{l_2}^{EE}] \cos 2(\varphi_{l_1} - \varphi_{l_2})$	$[\tilde{C}_{l_1}^{EE}(\mathbf{l} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{EE}(\mathbf{l} \cdot \mathbf{l}_2)] \cos 2(\varphi_{l_1} - \varphi_{l_2})$
$EB$	$[\tilde{C}_{l_1}^{EE} - \tilde{C}_{l_2}^{BB}] \sin 2(\varphi_{l_1} - \varphi_{l_2})$	$[\tilde{C}_{l_1}^{EE}(\mathbf{l} \cdot \mathbf{l}_1) - \tilde{C}_{l_2}^{BB}(\mathbf{l} \cdot \mathbf{l}_2)] \sin 2(\varphi_{l_1} - \varphi_{l_2})$
$BB$	$[\tilde{C}_{l_1}^{BB} + \tilde{C}_{l_2}^{BB}] \cos 2(\varphi_{l_1} - \varphi_{l_2})$	$[\tilde{C}_{l_1}^{BB}(\mathbf{l} \cdot \mathbf{l}_1) + \tilde{C}_{l_2}^{BB}(\mathbf{l} \cdot \mathbf{l}_2)] \cos 2(\varphi_{l_1} - \varphi_{l_2})$

in the flat-sky limit where scalar fields such as the CMB temperature  $T$  and a complex field  $(\mathcal{S}_1 \pm i\mathcal{S}_2)(\hat{\mathbf{n}})$  of spin  $\pm s$  can be expanded in the Fourier basis as

$$T(\mathbf{l}) = \int d\hat{\mathbf{n}} T(\hat{\mathbf{n}}) e^{-i\mathbf{l} \cdot \hat{\mathbf{n}}}, \quad (4)$$

$$[\mathcal{S}_1 \pm i\mathcal{S}_2](\mathbf{l}) = (\pm 1)^s \int d\hat{\mathbf{n}} [\mathcal{S}_1(\hat{\mathbf{n}}) \pm i\mathcal{S}_2(\hat{\mathbf{n}})] e^{\mp i s \varphi_1} e^{-i\mathbf{l} \cdot \hat{\mathbf{n}}},$$

where  $\varphi_1 = \cos^{-1}(\hat{\mathbf{n}} \cdot \hat{\mathbf{l}})$ . The complex field  $(Q \pm iU)(\hat{\mathbf{n}})$  is a spin  $\pm 2$  field, whose Fourier harmonics are referred as  $(E \pm iB)(\mathbf{l})$ .

Since the differential optical depth fluctuation is already constrained to be small, we work out the effects to first order in  $\delta\tau(\hat{\mathbf{n}})$ . We use  $\tau(\hat{\mathbf{n}})$  rather than  $\delta\tau(\hat{\mathbf{n}})$  to specify the fluctuations of optical depth for short. It is simple to show that patchy reionization induces modulations in observed CMB fields that is proportional to  $\tau(\hat{\mathbf{n}})$  to the first order. The effect of such a modulation is to correlate different CMB modes in Fourier space. The correlations can be compactly written as

$$\langle X(\mathbf{l}_1)X'(\mathbf{l}_2) \rangle_{\text{CMB}} = f_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2) \tau(\mathbf{l}), \quad (5)$$

where  $X, X' = T, E, B$ ,  $\mathbf{l} = \mathbf{l}_1 + \mathbf{l}_2$ ,  $f_{XX'}^\tau(\mathbf{l}, \mathbf{l}')$  is given in Table II, and  $\langle \dots \rangle_{\text{CMB}}$  signifies an ensemble average over CMB realizations with fixed  $\tau(\hat{\mathbf{n}})$  field.

The presence of  $\tau(\hat{\mathbf{n}})$  field breaks the rotational symmetry of the CMB field, correlating different modes which are not correlated assuming a Gaussian CMB field. Following [24], we construct a minimum variance quadratic estimator  $\hat{\tau}_{XX'}(\hat{\mathbf{n}})$  for  $\tau(\hat{\mathbf{n}})$  field, or  $\hat{\tau}_{XX'}(\mathbf{l})$  for  $\tau(\mathbf{l})$  in Fourier space.

$$\hat{\tau}_{XX'}(\mathbf{l}) = N_{XX'}(l) \int \frac{d^2 l_1}{(2\pi)^2} [X(\mathbf{l}_1)X'(\mathbf{l}_2)] F_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2),$$

where  $\mathbf{l} = \mathbf{l}_2 + \mathbf{l}_1$  and

$$N_{XX'}(l) = \left[ \int \frac{d^2 l_1}{(2\pi)^2} f_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2) F_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2) \right]^{-1}. \quad (6)$$

We derive the optimal  $F_{XX'}$  by minimizing the variance of  $\langle \hat{\tau}_{XX'}(\mathbf{l}) \hat{\tau}_{XX'}(\mathbf{l}') \rangle$ . For  $XX' = EE, BB$ , and  $TT$ ,

$$F_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2) = \frac{f_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2)}{2C_{l_1}^{XX,t} C_{l_2}^{XX,t}}. \quad (7)$$

For  $XX' = TB$  and  $EB$ ,

$$F_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2) = \frac{f_{XX'}^\tau(\mathbf{l}_1, \mathbf{l}_2)}{C_{l_1}^{XX,t} C_{l_2}^{XX',t}}, \quad (8)$$

where

$$C_l^{XX',t} = \tilde{C}_l^{XX'} + C_l^{XX',n}, \quad (9)$$

and  $C_l^{XX',n}$  is the noise power spectrum. We assume the detector noise is Gaussian and isotropic, to be known a priori. Furthermore, we assume a symmetric Gaussian instrumental beam so that the noise power spectrum is

$$C_l^{XX,n} = \Delta_X^2 e^{l^2 \Theta_{\text{fwhm}}^2 / (8 \ln 2)}, \quad (10)$$

where  $\Delta_X$  is the instrument noise for temperature ( $X = T$ ) or polarization ( $X = E, B$ ), and  $\Theta_{\text{fwhm}}$  is the full-width half-maximum (FWHM) of the Gaussian beam. We assume a fully polarized detector, for which  $\Delta_{E,B} = \sqrt{2} \Delta_T$ .

The variance of the minimum variance quadratic estimator is

$$\langle \tilde{\tau}_{XX'}(\mathbf{l}_1) \tilde{\tau}_{XX'}(\mathbf{l}_2) \rangle = (2\pi)^2 \delta(\mathbf{l}_1 + \mathbf{l}_2) \{ C_l^{\tau\tau} + N_{XX'}(l) \}, \quad (11)$$

where  $N_{XX'}(l)$  gives the dominant contribution to the variance for the  $EB$  and  $TB$  estimators.

#### IV. LENSING CONTAMINATION IN $\tau(\hat{\mathbf{n}})$ RECONSTRUCTION

The optical depth estimators described in the previous section neglect the effect of CMB lensing. In reality, both the CMB temperature and polarization fields are gravitationally lensed by inhomogeneities in the matter distribution between the last scattering surface and  $z = 0$ . In



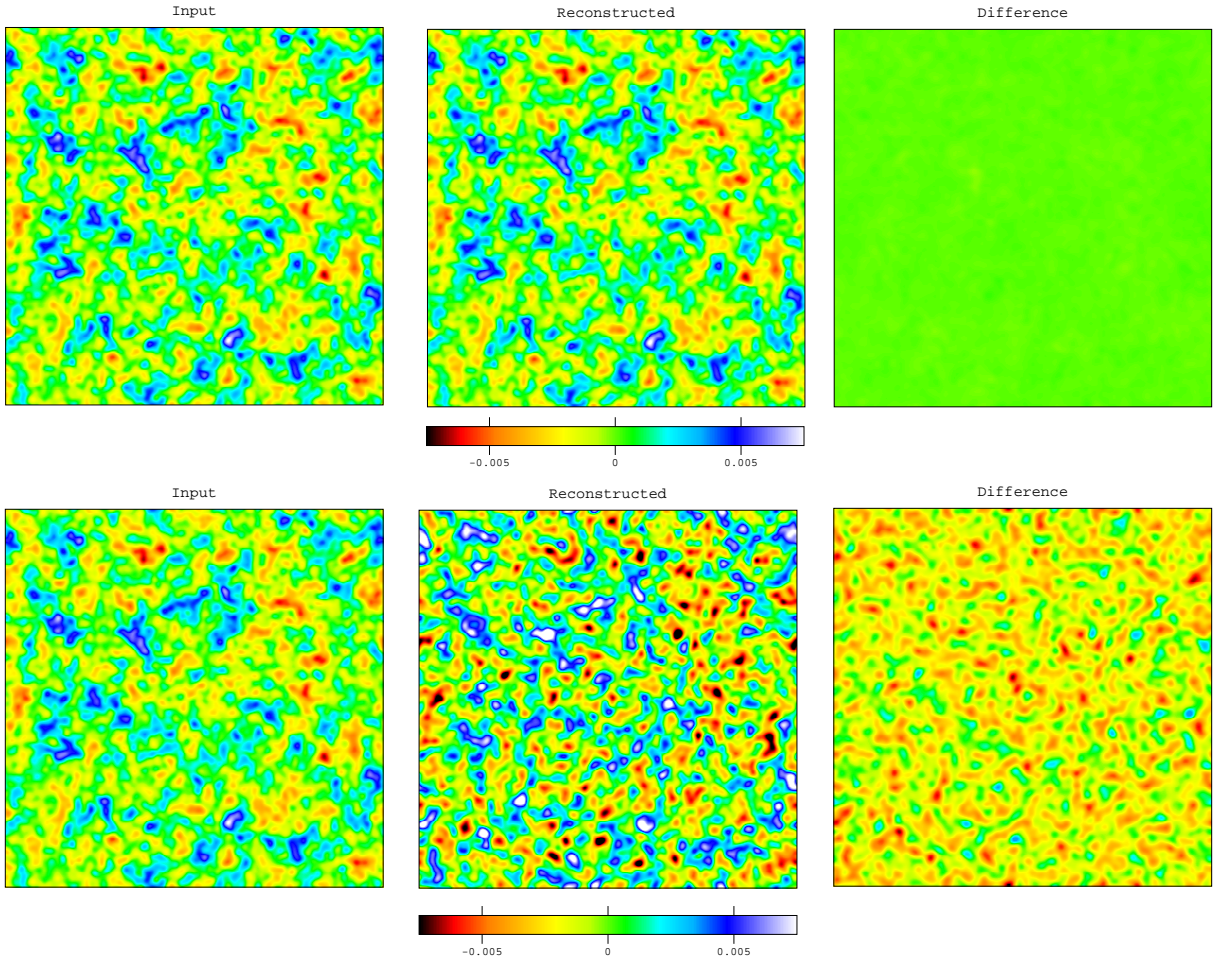


FIG. 2: Impact of lensing on the reconstruction of the optical depth fluctuation field  $\tau(\hat{\mathbf{n}})$ . The *upper panels* assume all the mode-coupling of CMB maps are generated by modulations of patchy reionization only. The upper left panel shows the input  $\tau(\hat{\mathbf{n}})$  map used to modulate the CMB fields, the upper middle panel shows the reconstructed  $\tau(\hat{\mathbf{n}})$  map from the CMB fields by applying the quadratic  $EB$  estimator, and the upper right panel shows the difference between the input/reconstructed  $\tau(\hat{\mathbf{n}})$  maps. The *lower panels* show the same quantities as the *upper panels* but with the lensing effect on the CMB maps included. As is clear from the lower middle and lower right panels, additional mode-coupling of CMB fields due to lensing contaminates the  $\tau(\hat{\mathbf{n}})$  reconstruction. In this plot, we consider the lensing signal to be only 2% of the fiducial value (i.e. the deflection angle power spectrum used here is  $C_l^{dd}/50$ ) to approximate the residual lensing signal after applying delensing procedure on the observed CMB maps. The reconstructed maps were averaged over 1000 CMB realizations for a fixed optical depth fluctuation  $\tau(\hat{\mathbf{n}})$  field. Each map is  $6 \times 6$  square degrees.

this section, we show that lensing significantly bias the  $\tau(\hat{\mathbf{n}})$  reconstruction.

Both the  $\tau(\hat{\mathbf{n}})$  field and the projected lensing potential  $\phi^{len}(\hat{\mathbf{n}})$  can generate non-Gaussianity by mixing modes and break the rotational invariance. This effect can be detected statistically by searching for the characteristic four point correlations. If the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator derived in the previous section were applied to the lensed CMB maps, it would also pick up significant spurious signal produced by lensing.

We now quantitatively calculate the lensing bias to the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator. Lensing simply deflects the path of CMB photons from the last scattering surface resulting in a remapping of the CMB temperature/polarization pat-

tern on the sky. The deflection angle  $\mathbf{d}(\hat{\mathbf{n}})$  is related to  $\phi^{len}(\hat{\mathbf{n}})$ , the lensing gravitational potential as

$$\mathbf{d}(\hat{\mathbf{n}}) = \nabla \phi^{len}(\hat{\mathbf{n}}). \quad (12)$$

The lensing potential  $\phi^{len}(\hat{\mathbf{n}})$  is given by

$$\phi^{len}(\hat{\mathbf{n}}) = -2 \int_0^{r_0} dr \frac{d_A(r_0 - r)}{d_A(r)d_A(r_0)} \Phi(r, \hat{\mathbf{n}}), \quad (13)$$

where  $d_A$  is the co-moving distance along the line of sight;  $r_0$  is the comoving distance to the surface of last scattering, and  $\Phi$  is gravitational potential [27].

Similar to the effect of screening from patchy reionization, a lensing potential mode with wavevector  $\mathbf{l}$  mixes the two polarization modes of wavevectors  $\mathbf{l}_1$  and  $\mathbf{l}_2 =$

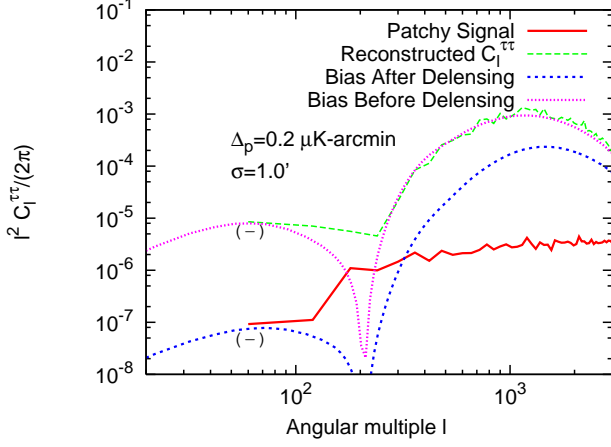


FIG. 3: Comparison of the biased reconstruction of patchy reionization power spectrum  $\hat{C}_l^{\tau\tau}$  with the input fiducial patchy reionization induced power spectrum  $C_l^{\tau\tau}$ . The solid red line shows the  $C_l^{\tau\tau}$  of the reionization scenario we chose. The dot magenta curve shows the biased estimate for  $C_l^{\tau\tau}$ , which was calculated using Eq. (15). The lensing bias is negative at low  $l$  but positive at high  $l$ . The minimum of the bias at  $l \sim 200$  is where the sign changes. The green dashed curve is the same quantity extracted from simulations after averaging over 1000 realizations. Our analytic expression matches well with the simulation. The blue short-dashed curve shows the reconstructed  $C_l^{\tau\tau}$  after delensing is applied. The delensing was performed using the quadratic minimum variance estimator for projected lensing potential.

$l - l_1$ . Taking the ensemble average of the CMB fields for the fixed  $\phi(\hat{\mathbf{n}})$  field, similar to Eq. (5), one gets

$$\langle \tilde{X}(\mathbf{l}_1) \tilde{X}'(\mathbf{l}_2) \rangle_{\text{CMB}} = f_{XX'}^{\phi}(\mathbf{l}_1, \mathbf{l}_2) \phi^{len}(\mathbf{l}). \quad (14)$$

The form of filters  $f_{XX'}^{len}(\mathbf{l})$  for different combinations of CMB field  $X$  and  $X'$  are given in Table II. The major difference between the filters of lensing potential estimator  $\hat{\phi}(\hat{\mathbf{n}})$  and those for the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator is the additional factors of  $\sim l^2$  that appear in those for lensing owing to the differential in Eq. (12). This differential nature of lensing significantly suppress the level of lensing estimator noise since  $l \gg 1$  and  $N_{XX'}^{lens}(l)$  is approximately proportional to  $l^{-2}$  (see Eq. (6)).

Let us consider a CMB sky that has been modified by both inhomogeneous reionization and lensing. Suppose we want to reconstruct  $\tau(\hat{\mathbf{n}})$  assuming that all of the non-Gaussianity is from patchy reionization, which is equivalent to applying the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator with filters designed to optimally reconstruct  $\tau(\hat{\mathbf{n}})$ . In this case, the estimator measures

$$\begin{aligned} \langle \hat{\tau}_{XX'}(\mathbf{l}) \rangle_{\text{CMB}} &= \tau(\mathbf{l}) \\ &+ \underbrace{N_{XX'}(l) \int \frac{d^2 l_1}{(2\pi)^2} f_{XX'}^{\phi} F_{XX'}^{\tau}(\mathbf{l}_1, \mathbf{l}_2) \phi^{len}(\mathbf{l})}_{\text{bias}}, \end{aligned} \quad (15)$$

where  $N_{XX'}$  is given by Eq. (6). The first term on the right hand side is the desired signal, and the second term

is a bias that owes to lensing. Note that  $f_{XX'}^{\phi}$  is the lensing filter (see Table II) and  $F_{XX'}^{\tau}(\mathbf{l}_1, \mathbf{l}_2)$  is given by Eq. (7) and (8).

We simulate a patchy reionization induced  $\tau(\hat{\mathbf{n}})$  field (model B in Table I) and modulate the CMB fields by the  $\tau(\hat{\mathbf{n}})$  field accordingly. We compare the reconstructed  $\tau(\hat{\mathbf{n}})$  with the input field in Fig. 2 (see Appendix B for details of the simulations). The  $\hat{\tau}(\hat{\mathbf{n}})$  estimator is unbiased if primordial CMB fields were unlensed and only affected by patchy reionization. However, in the presence of lensing, the reconstructed  $\tau(\hat{\mathbf{n}})$  deviates significantly from the fiducial signal. The lensing induced non-Gaussianity is roughly an order of magnitude larger than the patchy reionization induced non-Gaussianity.

The lensing induced non-Gaussianity could be reduced by applying lensing estimator  $\hat{\phi}(\hat{\mathbf{n}})$  to reconstruct the lensing potential, and then “remap” the observed CMB photons given the reconstructed  $\phi(\hat{\mathbf{n}})$  and Eq. (12). This process of subtracting the lensing effect from CMB is referred to as “delensing” (see Ref [35] for a review). To investigate the lensing bias after applying this delensing procedure, we assume the residual lensing potential power spectrum is only 2% of the fiducial value. The delensing fraction taken here is smaller than the predicted delensing fraction for future CMB experiment, using lensing maps either externally reconstructed from large scale structure/CMB temperature or from CMB polarization itself [36]. We find that even after delensing on the CMB map, the reconstructed  $\tau(\hat{\mathbf{n}})$  field is still significantly contaminated by the residual lensing signal, as shown in Fig. 2.

In Fig. 3, we show the reconstruction of optical depth fluctuation power spectrum  $\hat{C}_l^{\tau\tau}$ , and compare with the input power spectrum  $C_l^{\tau\tau}$ . Again we choose Model B for the reionization simulations, which has the highest level of  $\tau(\hat{\mathbf{n}})$  fluctuations. We find that the lensing induced spurious signal dominates over the fiducial signal by  $\sim \mathcal{O}(10 - 100)$ , especially for  $l \gtrsim 200$ . The theoretical prediction for the spurious patchy reionization signal from lensing which is calculated by Eq. (15), matches well with  $\hat{C}_l^{\tau\tau}$  from the simulation. Finally we show that even after applying the delensing procedure with lensing quadratic estimator [24], the reconstructed  $\hat{C}_l^{\tau\tau}$  is still biased by a factor of  $\sim 10$ . As we show in Fig. 3, the lensing induced  $C_l^{\tau\tau}$  has two bumps one peak at large scale  $l \sim 50$  and the other peaks at small scale  $l \sim 1000$ . It is caused by the lensing bias given by Eq. (15) is negative at low  $l$  and positive at high  $l$  with a transition at  $l \approx 200$ . This sign change is because the lensing bias involves the product of the lensing and tau filters [see Eq. (15)]. The product contains a mode coupling term  $\mathbf{l} \cdot \mathbf{l}_1$  which is caused by the derivative nature of lensing and gives the negative contribution at low  $l$ . Physically, the lensing of CMB does not generate new power in the CMB fluctuations, it only move power from large scale to small scales [27]. We note that in principle lensing reconstruction is also biased by the patchy reionization induced non-Gaussianity, however since lensing signal is

much larger than the patchy reionization signal, we don't expect a significant contamination from patchy reionization to lensing estimation.

## V. RECONSTRUCTING PATCHY REIONIZATION

### A. Unbiased Estimator

This section constructs an unbiased estimator for  $\tau(\hat{\mathbf{n}})$ . As with the quadratic estimator discussed in the previous section, among all the six estimators, the  $EB$  estimator has the highest  $S/N$  ratio, thus we focus on  $EB$  estimator in this section. For each multipole  $l$  we can define a 2-by-2 Fisher matrix  $F_l^{\alpha\beta}$ ,

$$F_l^{\alpha\beta} = \int \frac{d^2 l_1}{(2\pi)^2} f_{EB}^{(\alpha)}(l_1, l_2) (\mathbf{C}^{-1})_{l_1}^{EE} f_{EB}^{(\beta)}(l_1, l_2) (\mathbf{C}^{-1})_{l_2}^{BB}$$

where  $\alpha$  and  $\beta$  run over  $\tau$  and  $\phi$ . The element  $(F_l^{(-1)})^{\alpha\beta}$  of the inverse of Fisher matrix gives the variance of  $C_l^{\alpha\beta}$ . Hence the variance of  $C_l^{\tau\tau}$  is:

$$N(l) \equiv (\mathbf{F}^{-1})_l^{\tau\tau} = \frac{F_l^{\phi\phi}}{F_l^{\tau\tau} F_l^{\phi\phi} - (F_l^{\tau\phi})^2}. \quad (16)$$

This is the Gaussian noise term of the unbiased  $\hat{\tau}(\hat{\mathbf{n}})$  estimator, which we will use to calculate  $(S/N)^2$  in Fig. 4.

Starting from biased estimator  $\hat{\tau}(\hat{\mathbf{n}})$  and  $\hat{\phi}(\hat{\mathbf{n}})$  [24], we have

$$\begin{aligned} \langle \hat{\tau}(\mathbf{l}) \rangle_{\text{CMB}} &= \tau(\mathbf{l}) + \frac{F_l^{\tau\phi}}{F_l^{\tau\tau}} \phi(\mathbf{l}), \\ \langle \hat{\phi}(\mathbf{l}) \rangle_{\text{CMB}} &= \phi(\mathbf{l}) + \frac{F_l^{\phi\tau}}{F_l^{\phi\phi}} \tau(\mathbf{l}). \end{aligned} \quad (17)$$

One can then solve above equations for  $\tau(\mathbf{l})$  and  $\phi(\mathbf{l})$

$$\begin{aligned} \tau(\mathbf{l}) &= \frac{F_l^{\phi\phi} F_l^{\tau\tau} \langle \hat{\tau}(\mathbf{l}) \rangle_{\text{CMB}} - F_l^{\tau\phi} F_l^{\phi\phi} \langle \hat{\phi}(\mathbf{l}) \rangle_{\text{CMB}}}{F_l^{\tau\tau} F_l^{\phi\phi} - (F_l^{\tau\phi})^2}, \\ \phi(\mathbf{l}) &= \frac{F_l^{\phi\phi} F_l^{\tau\tau} \langle \hat{\phi}(\mathbf{l}) \rangle_{\text{CMB}} - F_l^{\tau\phi} F_l^{\tau\tau} \langle \hat{\tau}(\mathbf{l}) \rangle_{\text{CMB}}}{F_l^{\tau\tau} F_l^{\phi\phi} - (F_l^{\tau\phi})^2}. \end{aligned} \quad (18)$$

This estimator although unbiased is not a minimum variance estimator. In next subsection we compare the variance (Gaussian noise) of the minimum variance quadratic estimator with the variance of the unbiased estimator and show that there is only a marginal increase in the variance of the unbiased estimator in comparison to the variance of the minimum-variance estimator.

*Maximum Likelihood Estimator:* Given that the  $B$ -mode polarization is well mapped, Hirata and Seljak [26] found that for lensing reconstruction the maximum-likelihood estimator (which reduces the estimator noise from lensing) allows significantly better  $(S/N)^2$  than the quadratic estimator.

Following Ref. [26], the lensing maximum-likelihood estimator can be generalized to construct a unbiased maximum-likelihood estimator for  $\tau(\hat{\mathbf{n}})$  in the presence of lensing. The variance of the maximum-likelihood estimator for  $\tau(\hat{\mathbf{n}})$  is the same as that for the quadratic estimator  $\hat{\tau}(\hat{\mathbf{n}})$  with one exception— for maximum-likelihood estimator the denominator of Eq. (7) and Eq. (8) contains the unlensed CMB power spectrum, whereas the quadratic estimator noise contains the lensed CMB power spectrum [26]. The estimator noise of  $\tau(\hat{\mathbf{n}})$  reconstruction would no longer be saturated because of the lensed CMB power spectrum. Conceptually, the lensing or patchy reionization induced B-modes can be iteratively cleaned from the map, therefore we are able to reduce the post-cleaning B-mode power spectrum and thus reducing the noise in the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator. Our fundamental ability to clean the map is bounded by the sum of the unlensed CMB B-modes and the instrumental noise.

### B. Forecasting the Detectability of Patchy Reionization

The signal-to-noise for the detection of patchy reionization signal can be written as

$$\left(\frac{S}{N}\right)^2 = \left[ \frac{f_{sky}}{2} \sum_l (2l+1) \left( \frac{C_l^{\tau\tau}}{N(l)} \right)^2 \right], \quad (19)$$

where  $f_{sky}$  is the sky fraction;  $C_l^{\tau\tau}$  is the fiducial patchy reionization power spectrum, and  $N(l)$  is the leading order Gaussian noise of an estimator, given by Eq. (16) for the unbiased quadratic estimator and given by Eq. (6) for the biased minimum variance quadratic estimator.

In Fig. 4, dashed-lines (the lower two curves which almost overlap) compare the  $(S/N)^2$  of the biased and unbiased quadratic estimators. The left panel is for the CMBPol like experiment with noise  $\Delta_T = 1\mu\text{k-arcmin}$  and beam FWHM  $\Theta_{\text{fwhm}} = 4$  arcmin. The right panel is for the reference experiment with noise  $\Delta_T = 0.2\mu\text{k-arcmin}$  and beam FWHM  $\Theta_{\text{fwhm}} = 1$  arcmin. As is clear from figure the  $(S/N)^2$  of unbiased estimator is only slightly lower than the  $(S/N)^2$  of biased estimator for both CMBPol like experiment and the reference experiment. In another word, the variance of the unbiased estimator is only marginally more (percent-level) than the variance of the minimum-variance quadratic estimator. The reason for this is easy to understand—the contribution to the variance from the spurious  $\tau(\hat{\mathbf{n}})$  signal produced by lensing is much smaller than the intrinsic  $\hat{\tau}(\hat{\mathbf{n}})$  estimator noise.

In Fig. 4, dotted-lines (the upper two lines which almost overlap) compare the  $(S/N)^2$  of the biased/unbiased maximum-likelihood  $\hat{\tau}(\mathbf{l})$  estimators. For CMBPol-like experiment, the maximum-likelihood estimator can get  $(S/N)^2$  about a factor of 10 higher than the quadratic estimator.

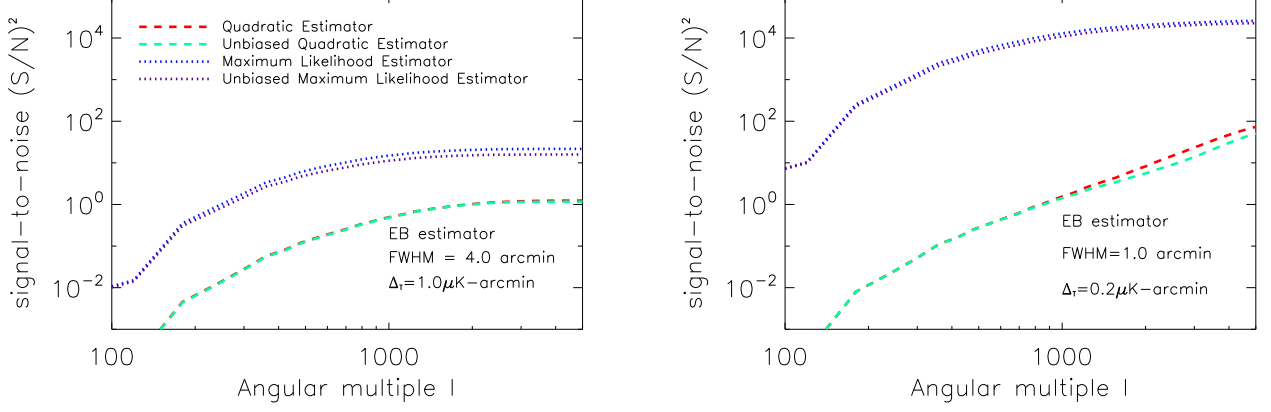


FIG. 4: The cumulative  $(S/N)^2$  for the  $EB$  estimator as a function of maximum multipole  $l$ . The left panel is for a CMBPol-like experiment with a beam of  $\Theta_{\text{FWHM}} = 4'$  and the noise sensitivity  $\Delta_p = 1\mu\text{K-arcmin}$ . The right panel is a more sensitive experiment with  $\Theta_{\text{FWHM}} = 1'$  and  $\Delta_p = 0.2\mu\text{K-arcmin}$ . In both panels, we use the optical depth power spectrum  $C_l^{\tau\tau}$  of model B in Fig.1. The green/red dotted curves (lower two curves) which almost overlap are the  $(S/N)^2$  of the biased/unbiased quadratic estimators, whose noise level is given by Eq. (6) and (16) respectively. The magenta/blue dashed curves (the top two curves that nearly overlap), are the  $(S/N)^2$  of the biased/unbiased maximum-likelihood estimators. The maximum-likelihood estimator is calculated from the minimum variance quadratic estimator except with the lensed CMB power spectrum replaced by the primary CMB power-spectrum (without the presence of lensing or patchy reionization effects) as suggested in [26].

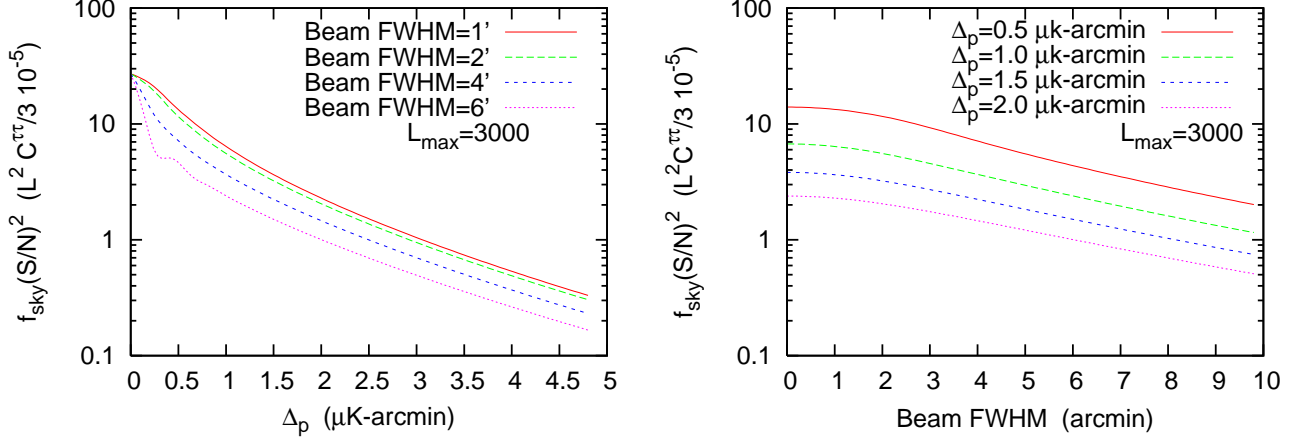


FIG. 5: Dependence of the total  $(S/N)^2$  on instrumental sensitivity,  $\Delta_p$ , (left panel) and beam size,  $\Theta_{\text{FWHM}}$  (right panel). We plot  $EB$  unbiased quadratic estimator as an example and calculate cumulative  $(S/N)^2$  up to  $l_{\text{max}} = 3000$ . *Left panel*, we fix beam size with 1, 2, 4, and 6 arcmin respectively to calculate the  $(S/N)^2$  dependence on instrumental sensitivity. *Right panel*, we fix  $\Delta_p$  with 0.5, 1.0, 1.5, and  $2.0\mu\text{K-arcmin}$  to show the beam size dependence of  $(S/N)^2$ . The approximate fitting functions are given in Eq. (20).

In Fig. 5, we show the total  $(S/N)^2$  from the unbiased quadratic estimator as a function of instrumental beam size and detector sensitivity respectively. We find that the  $(S/N)$  for a constant  $l^2 C_l^{\tau\tau}/(2\pi)$  and for experiment with  $\Delta_p > 1\mu\text{K-arcmin}$  can be approximated as

$$\left(\frac{S}{N}\right)^2 \approx 5f_{\text{sky}} \left(\frac{l^2 C_l^{\tau\tau}/2\pi}{5 \times 10^{-6}}\right) \exp\left(\frac{-0.56\Delta_p}{1\mu\text{K-arcmin}}\right) \times \exp\left(\frac{-0.09\Theta_{\text{FWHM}}}{1'}\right). \quad (20)$$

The  $(S/N)^2$  is more sensitive to the instrumental sensitivity rather than the beam size. For a reference patchy reionization signal  $l^2 C_l^{\tau\tau}/2\pi = 5 \times 10^{-6}$ , for a CMBPol-like or CoRE-like [37] experiment we expect  $(S/N)^2 \sim \mathcal{O}(1)$ . For the future ground-based experiments such as the POLAR Array with  $\Delta_p = 1.41\mu\text{K-arcmin}$ ,  $\Theta_{\text{FWHM}} = 1'$ , and sky coverage  $100 \text{ deg}^2$ , we expect  $(S/N)^2 \sim \mathcal{O}(0.01)$ .



## VI. SUMMARY AND DISCUSSION

Reionization marks the epoch in which the vast majority of the hydrogen in the Universe was ionized since cosmological recombination. When and how reionization occurred is at present poorly constrained. In addition to pinning down the epoch of this cosmic phase transition, constraints on the reionization history provides us information about the formation of early galaxies. Inhomogeneous reionization would have generated fluctuations in the Thomson scattering optical depth  $\tau(\hat{\mathbf{n}})$  among different lines of sights at the level  $\sim 10^{-3}$ . These modulations would modify the primordial CMB temperature and polarization anisotropies by inducing a directionally dependent screening. Such screening couples different modes of CMB, converting  $E$ -modes to  $B$ -modes, and introduces non-Gaussian signals.

In this paper, we used a technique that exploits the non-Gaussianities in the CMB sourced by reionization to study this process, as first proposed in Dvorkin and Smith [12]. We have introduced the minimum variance quadratic estimator in an intuitive flat sky limit and compared it with the estimator for lensing potential reconstruction [24]. Lensing induced non-Gaussian features would produce a spurious  $\tau(\hat{\mathbf{n}})$  signal that is at least an order of magnitude higher than our semi-analytical models predict from patchy reionization. We showed that ignoring the lensing contamination would significantly bias the reconstruction of optical depth fluctuation field  $\tau(\hat{\mathbf{n}})$ . Even after applying a delensing procedure that used the minimum variance quadratic estimator for the lensing potential  $\phi(\hat{\mathbf{n}})$ , the residual lensing bias on the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator was still comparable with the fiducial value. As a solution, we constructed an unbiased estimator to simultaneously reconstruct  $\tau(\hat{\mathbf{n}})$  and the lensing potential  $\phi(\hat{\mathbf{n}})$  such that the estimate of  $\tau(\hat{\mathbf{n}})$  is not biased by lensing. We found that the  $S/N$  of the unbiased estimator is only degraded at the percent level compared to the original biased  $\tau(\hat{\mathbf{n}})$  estimator.

We studied the detectability of patchy reionization by considering more detailed  $\tau(\hat{\mathbf{n}})$  fields using semi-numerical reionization models, which unfortunately yield an order-of-magnitude smaller signal than previously considered [12]. As a result, we found that with the unbiased estimator, a CMBPol-like experiment could achieve a marginal detection of patchy reionization with  $(S/N)^2 \sim 1 - 10$ . We characterized the estimator noise for various instrumental properties. We find that the  $S/N$  is only weakly sensitive to the FWHM of detector beam with a factor of  $\sim 2$  degradation of  $(S/N)^2$  by increase FWHM from  $1'$  to  $6'$ . While the  $(S/N)^2$  decreases by a factor of  $\sim 2$  by increase instrumental noise from  $0.5$  to  $2 \mu\text{k-arcmin}$ . Similar scaling with instrumental characteristics have been quantified for lensing reconstruction in [24].

Large scale CMB fields are also modulated by smaller scale  $\tau(\hat{\mathbf{n}})$  fluctuations due to patchy reionization. As we construct the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator in flat sky limit, we ignore

the patchy reionization signal from large scale  $E/B$ -mode which is generated via scattering of the local CMB temperature quadrupole by ionized bubbles. The  $S/N$  will be increased by a factor of  $\sim 2$  by considering such signal on large scales [12].

Although the predicted  $S/N$  for a patchy reionization detection is only marginal for a CMBPol-like experiment, one can cross-correlate with other cosmological data sets that are sensitive to the properties of patchy reionization. The same population of ionized bubbles would not only induce line of sight dependent optical depth of CMB, but also correlate with the distribution of galaxies or the redshifted 21cm signal.

At large scales, it is expected that the distribution of galaxies correlates well with the neutral gas distribution [28]. One can estimate the  $(S/N)_\tau$  for a patchy reionization detection as

$$(S/N)_\tau^2 = l_{\text{char}}^2 (C_{\text{char}}^{\tau\tau})^2 / (C_l^{\hat{\tau}\hat{\tau}})^2 \quad (21)$$

where  $l_{\text{char}}$  is the characteristic multipole that contributes to the  $S/N$  ( $l_{\text{char}} \sim 10^3$ ), and  $C_l^{\hat{\tau}\hat{\tau}}$  is the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator variance. The scaling factor  $l_{\text{char}}^2$  is an estimate for the number of modes that are contributing to the signal (the result really does not rely on the fraction of the sky  $\tau(\hat{\mathbf{n}})$  is estimated).

An estimate for signal-to-noise that can be obtained in cross correlation  $(S/N)_{\tau g}$  is

$$\begin{aligned} (S/N)_{\tau g}^2 &\approx l_{\text{char}}^2 f_{\text{sky}} f_{\text{reion}} (C_{\text{char}}^{\tau g})^2 / (C_l^{\hat{\tau}\hat{\tau}} C_{gg}) \\ &= l_{\text{char}}^2 f_{\text{sky}} f_{\text{reion}} r^2 C_{\text{char}}^{\tau\tau} / C_l^{\hat{\tau}\hat{\tau}} \\ &= l_{\text{char}} f_{\text{sky}} f_{\text{reion}} r^2 (S/N)_\tau, \end{aligned} \quad (22)$$

where  $f_{\text{sky}}$  and  $f_{\text{reion}}$  are the fraction of the sky and reionization over which surveys overlap,  $r$  is the cross correlation coefficient of galaxies and the  $\tau(\hat{\mathbf{n}})$  field over the same projected volume as the galaxy survey ( $r \sim 1$  on large scale).

Noting that  $f_{\text{sky}} \lesssim 10^{-4}$  is the current size for  $z \sim 7$  galaxy surveys, it would take a very ambitious survey to enhance the  $\tau$  signal in cross correlation compared to in the auto-power. Correlating with the diffuse background light from early galaxies – the cosmic infrared background – is a related and intriguing possibility since then  $f_{\text{sky}} f_{\text{reion}} \sim 1$  (although, lower redshift emission may be a significant noise source in this case) and may deserve further study.

The final possibility that we discuss is cross correlating with a survey of redshifted 21cm emission from intergalactic neutral hydrogen. Such surveys do span a significant fraction of the sky and the first generation of such endeavors will be in a noise-dominated regime in which they could benefit from cross-correlation [38] (Note that cross correlating with  $\tau$  would be of little interest if there existed high  $S/N$  21cm maps). However, redshifted 21cm analyses remove the modes with small line-of-sight

projected wavevectors in the act of foreground cleaning, which unfortunately are the modes that contribute to the  $\tau$  signal [39]. Thus, there would be little signal in this cross correlation.

### Acknowledgments

We thank C. Dvorkin and K. M. Smith for helpful dis-

cussions. APSY gratefully acknowledges support from IBM Einstein fellowship and funding from NASA award number NNX08AG40G and NSF grant number AST-0807444. MM is supported by the NASA Einstein fellowship. JY is supported by the SNF Ambizione grant. MZ is supported by the National Science Foundation under PHY-0855425 and AST-0907969, and by the David and Lucile Packard Foundation and the John D. and Catherine T. MacArthur Foundation.

- 
- [1] X. Fan et al., *Astron. J.* **132**, 117 (2006), [astro-ph/0512082](#).
  - [2] E. Komatsu, K. M. Smith, J. Dunkley, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. R. Nolte, L. Page, et al., *Astrophysical Journal Supplement* **192**, 18 (2011), 1001.4538.
  - [3] R. Barkana and A. Loeb, *Phys. Rep.* **349**, 125 (2001), [arXiv:astro-ph/0010468](#).
  - [4] S. R. Furlanetto, M. Zaldarriaga, and L. Hernquist, *Astrophys. J.* **613**, 1 (2004), [arXiv:astro-ph/0403697](#).
  - [5] M. McQuinn, A. Lidz, O. Zahn, S. Dutta, L. Hernquist, and M. Zaldarriaga, *Mon. Not. R. Astron. Soc.* **377**, 1043 (2007), [arXiv:astro-ph/0610094](#).
  - [6] J. Miralda-Escude, *Astrophys. J.* **501**, 15 (1998), [arXiv:astro-ph/9708253](#).
  - [7] M. McQuinn, L. Hernquist, M. Zaldarriaga, and S. Dutta, *Mon. Not. R. Astron. Soc.* **381**, 75 (2007), 0704.2239.
  - [8] M. Ouchi, K. Shimasaku, H. Furusawa, T. Saito, M. Yoshida, M. Akiyama, Y. Ono, T. Yamada, K. Ota, N. Kashikawa, et al., *Astrophys. J.* **723**, 869 (2010), 1007.2961.
  - [9] T. Totani, N. Kawai, G. Kosugi, K. Aoki, T. Yamada, M. Iye, K. Ohta, and T. Hattori, *PASJ* **58**, 485 (2006), [arXiv:astro-ph/0512154](#).
  - [10] M. McQuinn, A. Lidz, M. Zaldarriaga, L. Hernquist, and S. Dutta, *Mon. Not. R. Astron. Soc.* **388**, 1101 (2008), 0710.1018.
  - [11] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, *Phys. Rep.* **433**, 181 (2006), [arXiv:astro-ph/0608032](#).
  - [12] C. Dvorkin and K. M. Smith, *Phys. Rev. D* **79**, 043003 (2009), 0812.1566.
  - [13] R. A. Sunyaev and Y. B. Zeldovich, *Astrophys. Space Sci.* **7**, 20 (1970).
  - [14] R. A. Sunyaev and I. B. Zeldovich, *Ann. Rev. Astron. Astrophys.* **18**, 537 (1980).
  - [15] W. Hu, *Astrophys. J.* **529**, 12 (2000), [arXiv:astro-ph/9907103](#).
  - [16] M. McQuinn, S. R. Furlanetto, L. Hernquist, O. Zahn, and M. Zaldarriaga, *Astrophys. J.* **630**, 643 (2005), [arXiv:astro-ph/0504189](#).
  - [17] M. G. Santos, A. Cooray, Z. Haiman, L. Knox, and C.-P. Ma, *Astrophys. J.* **598**, 756 (2003), [arXiv:astro-ph/0305471](#).
  - [18] I. T. Iliev, U.-L. Pen, J. R. Bond, G. Mellema, and P. R. Shapiro, *Astrophys. J.* **660**, 933 (2007), [arXiv:astro-ph/0609592](#).
  - [19] P. Zhang, U.-L. Pen, and H. Trac, *Mon. Not. R. Astron. Soc.* **347**, 1224 (2004), [arXiv:astro-ph/0304534](#).
  - [20] A. Gruzinov and W. Hu, *Astrophys. J.* **508**, 435 (1998), [arXiv:astro-ph/9803188](#).
  - [21] O. Doré, G. Holder, M. Alvarez, I. T. Iliev, G. Mellema, U.-L. Pen, and P. R. Shapiro, *Phys. Rev. D* **76**, 043002 (2007), [arXiv:astro-ph/0701784](#).
  - [22] N. Aghanim, S. Majumdar, and J. Silk, *Reports on Progress in Physics* **71**, 066902 (2008), 0711.0518.
  - [23] G. P. Holder, K. M. Nollett, and A. van Engelen, *ArXiv e-prints* (2009), 0907.3919.
  - [24] W. Hu and T. Okamoto, *Astrophys. J.* **574**, 566 (2002).
  - [25] T. Okamoto and W. Hu, *Phys. Rev. D* **67**, 083002 (2003), [arXiv:astro-ph/0301031](#).
  - [26] C. M. Hirata and U. Seljak, *Phys. Rev. D* **68**, 083002 (2003), [arXiv:astro-ph/0306354](#).
  - [27] A. Lewis and A. Challinor, *Phys. Rep.* **429**, 1 (2006), [arXiv:astro-ph/0601594](#).
  - [28] O. Zahn, A. Lidz, M. McQuinn, S. Dutta, L. Hernquist, M. Zaldarriaga, and S. R. Furlanetto, *Astrophys. J.* **654**, 12 (2007), [arXiv:astro-ph/0604177](#).
  - [29] O. Zahn, A. Mesinger, M. McQuinn, H. Trac, R. Cen, and L. E. Hernquist, *ArXiv e-prints* (2010), 1003.3455.
  - [30] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, *Astrophys. J.* **379**, 440 (1991).
  - [31] C. Lacey and S. Cole, *Mon. Not. R. Astron. Soc.* **262**, 627 (1993).
  - [32] S. R. Furlanetto and S. P. Oh, *Mon. Not. R. Astron. Soc.* **363**, 1031 (2005), [arXiv:astro-ph/0505065](#).
  - [33] D. Tseliakhovich and C. Hirata, *Phys. Rev. D* **82**, 083520 (2010), 1005.2416.
  - [34] N. Dalal, U. Pen, and U. Seljak, *JCAP* **11**, 7 (2010), 1009.4704.
  - [35] K. M. Smith, A. Cooray, S. Das, O. Doré, D. Hanson, C. Hirata, M. Kaplinghat, B. Keating, M. LoVerde, N. Miller, et al., *ArXiv e-prints* (2008), 0811.3916.
  - [36] K. M. Smith, D. Hanson, M. LoVerde, C. M. Hirata, and O. Zahn, *ArXiv e-prints* (2010), 1010.0048.
  - [37] The CORe Collaboration, *ArXiv e-prints* (2011), 1102.2181.
  - [38] A. Lidz, S. R. Furlanetto, S. P. Oh, J. Aguirre, T.-C. Chang, O. Doré, and J. R. Pritchard, *ArXiv e-prints* (2011), 1104.4800.
  - [39] M. McQuinn, O. Zahn, M. Zaldarriaga, L. Hernquist, and S. R. Furlanetto, *Astrophys. J.* **653**, 815 (2006), [arXiv:astro-ph/0512263](#).
  - [40] W. Hu, S. DeDeo, and C. Vale, *New Journal of Physics* **9**, 441 (2007), [arXiv:astro-ph/0701276](#).
  - [41] J. Yoo and M. Zaldarriaga, *Phys. Rev. D* **78**, 083002 (2008), 0805.2155.
  - [42] J. Yoo, M. Zaldarriaga, and L. Hernquist, *Phys. Rev. D*

81, 123006 (2010), 1005.0847.

[43] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000), arXiv:astro-ph/9911177.

### Appendix A: Reconstructing Inhomogeneous Reionization $\tau(\hat{\mathbf{n}})$ : Simulation Pipeline

Our simulation pipeline of optical depth  $\tau(\hat{\mathbf{n}})$  field reconstruction follows the procedure in [40], and we modified the code developed for lensing reconstruction in [41, 42]. First, we generate primordial CMB polarization  $Q^{pri}(\hat{\mathbf{n}})$  and  $U^{pri}(\hat{\mathbf{n}})$  maps as Gaussian realizations of CMB power spectrum. We choose a standard fiducial model with a flat  $\Lambda$ CDM cosmology, with parameters given by  $\Omega_b = 0.045$ ,  $\Omega_c = 0.23$ ,  $H_0 = 70.5$ ,  $n_s = 0.96$ ,  $n_t = 0.0$ , and  $\tau = 0.08$ . We calculate the theoretical lensed and unlensed CMB power spectrum from publicly available code CAMB [43]. The primordial CMB polarizations maps are then transformed according to Eq. (3) to include the effect of patchy reionization. The  $\tau(\hat{\mathbf{n}})$  field was generated from a reionization simulation described in Section II.

To include the effect of lensing, we generate a realization of lensing deflection field  $\mathbf{d}(\hat{\mathbf{n}})$  and transform the CMB fields  $\tilde{Q}(\hat{\mathbf{n}})$  and  $\tilde{U}(\hat{\mathbf{n}})$  to  $Q(\hat{\mathbf{n}})$  and  $U(\hat{\mathbf{n}})$  according to

$$(Q \pm iU)(\hat{\mathbf{n}}) = (\tilde{Q} \pm i\tilde{U})(\hat{\mathbf{n}} + \mathbf{d}(\hat{\mathbf{n}})). \quad (\text{A1})$$

The deflection angle at each point  $\hat{\mathbf{n}}$  is calculated by taking the gradient of the lensing potential. The lensing potential power spectrum is generated using CAMB which was run with nonlinear corrections using halofit [43].

Since we want to quantify the lensing contamination, we have several pipelines with different level of lensing signal being removed. We define a de-lensing factor  $\alpha$ , as  $C_{\text{delen}, \phi\phi} = C_{\text{theory}, \phi\phi} / \alpha$ , where  $\alpha = 1$  correspond to no de-lensing,  $\alpha \rightarrow \infty$  corresponds to perfect delensing, we use  $\alpha = 50$  for Fig. (2).

We then Fourier transform CMB polarization maps to get  $E(\mathbf{l})$  and  $B(\mathbf{l})$  maps. Finally we multiply CMB  $E(\mathbf{l})$  and  $B(\mathbf{l})$  maps by Gaussian beam in Fourier space and add instrumental noise.

More specifically, we closely follow Hu et al [40] to re-write the  $\hat{\tau}(\hat{\mathbf{n}})$  estimator which is more efficient to evaluate computationally. We re-write the estimator in real space

$$\hat{\tau}_1^{EB} = -N_l^{EB} \int d^2\hat{\mathbf{n}} e^{-i\hat{\mathbf{n}} \cdot \mathbf{l}} \text{Re} \{ [\mathbf{G}^{EB}(\hat{\mathbf{n}}) L^{B*}(\hat{\mathbf{n}})] \}. \quad (\text{A2})$$

The field  $\mathbf{G}^{EB}$  is built from the observed  $E(\mathbf{l})$  field (including contributions from lensing and patchy reionization) as

$$\mathbf{G}_1^{EB} = \frac{C_l^{EE}}{(C_l^{EE} + N_l^{EE})} E(\mathbf{l}) e^{2i\varphi_1}. \quad (\text{A3})$$

and  $L^B$  is given by

$$L_1^B = \frac{B(\mathbf{l})}{(C_l^{BB} + N_l^{BB})} e^{2i\varphi_1}. \quad (\text{A4})$$

$N_l^{EB}$  is a normalization coefficient which ensure the unbiasedness of the estimator [40]. The results of our simulations are shown in Fig. (2).

### Appendix B: Unbiased Minimum Variance Quadratic Estimator for Patchy Reionization

This section discusses the quadratic estimator for the patchy reionization induced optical depth fluctuation field  $\tau(\hat{\mathbf{n}})$  in a more general context, demonstrating that the estimator  $\hat{\tau}(\hat{\mathbf{n}})$  used in the text is the minimum variance estimator in the limit that the signal-to-noise ratio in lensing estimator  $\hat{\phi}(\hat{\mathbf{n}})$  is much higher than in  $\hat{\tau}(\hat{\mathbf{n}})$ .

In Fourier space, the general unbiased quadratic estimator for  $\hat{\phi}_1$  and  $\hat{\tau}_1$  is

$$\begin{aligned} \hat{\tau}_1 &= \sum_{\mathbf{l}_i} Q_{\mathbf{l}_i} X(\mathbf{l} - \mathbf{l}_i) X'(\mathbf{l}_i) - \text{Tr}(Q f_{XX'}^\phi) \phi_1 \\ \hat{\phi}_1 &= \sum_{\mathbf{l}_i} P_{\mathbf{l}_i} X(\mathbf{l} - \mathbf{l}_i) X'(\mathbf{l}_i) - \text{Tr}(P f_{XX'}^\tau) \tau_1 \end{aligned} \quad (\text{B1})$$

where  $Q_{\mathbf{l}_i}$  and  $P_{\mathbf{l}_i}$  are some weighting functions,  $f_{XX'}^\tau$  and  $f_{XX'}^\phi$  is the same as in Eq. (5) and (14). The sum does not include  $\mathbf{l}_i = 0$  and we are using  $\text{Tr}(X)$  as shorthand for  $\sum_{\mathbf{l}_i} X_{\mathbf{l}_i}$ . Noting that  $\langle X(\mathbf{l} - \mathbf{l}_i) X'(\mathbf{l}_i) \rangle = f_{XX'}^\phi(\mathbf{l} - \mathbf{l}_i, \mathbf{l}_i) \phi_1 + f_{XX'}^\tau(\mathbf{l} - \mathbf{l}_i, \mathbf{l}_i) \tau_1$ , we can write the above equation as (if we substitute the unbiased estimator  $\hat{\phi}_1$  and  $\hat{\tau}_1$  on the R.H.S.)

$$\mathbf{A} \times \begin{pmatrix} \hat{\tau}_1 \\ \hat{\phi}_1 \end{pmatrix} = \begin{pmatrix} \sum_{\mathbf{l}_i} Q_{\mathbf{l}_i} X(\mathbf{l} - \mathbf{l}_i) X'(\mathbf{l}_i) \\ \sum_{\mathbf{l}_i} P_{\mathbf{l}_i} X(\mathbf{l} - \mathbf{l}_i) X'(\mathbf{l}_i) \end{pmatrix} \quad (\text{B2})$$

where

$$\mathbf{A} \equiv \begin{pmatrix} 1 & \text{Tr}(Q f_{XX'}^\phi) \\ \text{Tr}(P f_{XX'}^\tau) & 1 \end{pmatrix}. \quad (\text{B3})$$

Thus, the general unbiased quadratic estimator for  $\tau$  alone is

$$\hat{\tau}_1 = \sum_{\mathbf{l}_i} [[\mathbf{A}^{-1}]_{11} Q_{\mathbf{l}_i} + [\mathbf{A}^{-1}]_{12} P_{\mathbf{l}_i}] T(\mathbf{l} - \mathbf{l}_i) T(\mathbf{l}_i) \quad (\text{B4})$$

We want to derive the weighting functions  $Q_{\mathbf{l}_i}$  and  $P_{\mathbf{l}_i}$  that give us the unbiased minimum variance estimator  $\hat{\tau}_1$ . The estimator variance is

$$\begin{aligned} \text{var}[\hat{\tau}_1] &= \langle |\hat{\tau}_1|^2 \rangle \\ &= 2 \sum_{\mathbf{l}_i} X_{\mathbf{l}_i} C_{\mathbf{l}_i} X_{\mathbf{l}_i} C_{1-\mathbf{l}_i}, \end{aligned} \quad (\text{B5})$$

where

$$X_{\mathbf{l}_i} = [\mathbf{A}^{-1}]_{11} Q_{\mathbf{l}_i} + [\mathbf{A}^{-1}]_{12} P_{\mathbf{l}_i}. \quad (\text{B6})$$

To derive the minimum variance estimator, we want to minimize  $\text{var}[\hat{\tau}_1]$  subject to the conditions that  $\text{Tr}[P f_{XX'}^\tau] = 1$  and  $\text{Tr}[Q f_{XX'}^\phi] = 1$ . Rather than go through this exercise, let us note first that at relevant multipole  $f_{XX'}^\tau \ll f_{XX'}^\phi$ , because of the factor of  $\sim l^2$  that contributes to  $f_{XX'}^\phi$ . Let us also note that the weighting function  $P_i$  that is optimal for simultaneously estimating  $\phi_1$  with  $\tau_1$  should be nearly identical to the minimum variance quadratic weighting for estimating just  $\phi_1$  because  $\tau_1$  is a weak contaminant of lensing. Second note that  $[\mathbf{A}^{-1}]_{12} \propto \text{Tr}[P f_{XX'}^\tau] \ll 1$  (since  $\text{Tr}[P f_{XX'}^\phi] = 1$ ), and thus  $[\mathbf{A}^{-1}]_{12}$  has magnitude that is much less than that of  $[\mathbf{A}^{-1}]_{11}$ . Not only  $[\mathbf{A}^{-1}]_{11} \gg [\mathbf{A}^{-1}]_{12}$ , but note

the scaling  $Q_1/P_1 \sim l^2 \gg 1$ , therefore, we are justified in ignoring the second term in Eq. (B6) and one can show that the minimizing Eq. (B5) subject to the constraint  $\text{Tr}[Q f_{XX'}^\tau] = 1$  yields

$$Q_{1_i} = \text{Tr}[C^{-1} f_{XX'}^\tau C^{-1} f_{XX'}^\tau]^{-1} C_{1_i}^{-1} f_{XX'}^\tau (\mathbf{1}_i, \mathbf{1} - \mathbf{1}_i) C_{1-i}^{-1}, \quad (\text{B7})$$

which yields the identical estimator to that used in the text. Furthermore, because the variance of  $\hat{\tau}_1$  is dominated by  $Q_{1_i}$ , this explains why our unbiased estimator that accounts for  $\phi_1$  yields a result that is not much different than the biased minimum variance estimator.